Mini-Sentinel Methods:
Accomplishments and lessons learned
(with comments about the Vaccine Safety Datalink)

Jennifer Nelson, PhD
Sebastian Schneeweiss MD, ScD

June 3, 2011

Agenda

- Framework for post-marketing surveillance
- Vaccine Safety Datalink (VSD)
- Mini-Sentinel methods development to date
  - Data capacity
  - Distributed methods
  - Signal alerting strategies
- Needs and recommendations

Stages of post-marketing surveillance

Aim = Identify excess risk	All (suspected and unanticipated) adverse events (AEs), all products	Specific AE:product pairs of prior concern	A highly suspected AE:product pair
Approach	Consider many AEs or AE:product pairs (100's, 1000's)	Repeated monitoring or one-time expedited analysis of AE:product pairs (typically 5-10)	One-time, in-depth and rigorous investigation of a single pair
Example	Data mining of spontaneous reports	Active surveillance in Mini-Sentinel and VSD using coded electronic health information	Retrospective, formal epidemiological study using individual-level data, validated AEs, richer confounders

Sentinel Initiative Vision*

- System will be able to refine safety signals in near real-time. This will require the following capabilities:
  - rapidly defining exposed cohorts;
  - establishing algorithms to capture health outcomes of interest;
  - using sophisticated modular programs capable of running evaluations with minimal input from epidemiologists and clinicians and limited or no ad hoc programming; and
  - developing a framework to guide methodological approaches for safety surveillance evaluations that include confounding adjustment.
- Approaches for signal generation will be under development.

* Within the next 3 years

Map of methodologic domains

Data capacity	Distributed methods	Signal alerting
• **Integrity** – Common data model – Data completeness – Data validity – HOI validation • **Environments** – Claims – EHRs • Ambulatory • Inpatient – Registries – Other (blood banks, genetic data, etc.)	• **Distribution and retrieval** • **Anonymous linkage across sources** • **Distributed multivariable analysis** – Horizontal – Vertical	• **Design & validity** – Expedited design choice – Automated confounding adjustment • **Performance of** – Sequential testing – Non test-based – Decision analytic approaches • **Special aspects** – Drugs, vaccines, biologics, devices

Applications
• Oral antidiabetic agents and MI, rotavirus vaccine and intussusception, etc.

VSD: Data capacity

- **Integrity**
  - Quality of HMO vaccine database (Mullooly et al., AJE 1999)
  - Predictive value of seizure ICD-9 code (Shui et al., Vaccine 2009)
  - Accuracy of flu vaccine data in MCOs (Sy et al., Vaccine 2010)
- **Accessibility**
  - Rapid assessment of flu vaccine coverage (Lewis et al., MMWR 2005)
  - Active surveillance pilot to detect early signals (Davis et al., Epidemiol 2005)
  - VSD near real-time model & infrastructure (Baggs et al, Pediatrics 2011)
- **Diversification**
  - Enhancing detection w/EMRs (Hinrichsen et al., J Am Med Inform Assoc 2007)
  - Detecting vaccine AEs in clinical notes (Hazelhurst et. al. Vaccine 2009)

VSD: Distributed analysis methods & applications

- Developed a **prospective safety monitoring** framework
  - Real-time vaccine safety surveillance (Lieu et al., Med Care 2007)
  - Near real-time flu vaccine safety surveillance (Greene et al, AJE 2010)
- A system is emerging for **rapid signal evaluation**
  - Lessons learned to reduce false positives (Yih et al., Pediatrics 2011)
- Have evaluated these systems via **applications in practice**
  - MMRV and febrile seizures (Klein et al., Pediatrics 2010)
  - Rotavirus vaccine & intussusception (Belongia et al., Ped Inf Dis 2010)
  - Tdap safety in adolescents and adults (Yih et al, Vaccine 2009)
  - Ongoing: HPV, Pentacel, Kinrix, PCV13, others…

VSD: Alerting strategies

- Study **design**
  - Comparing designs for active surveillance (McClure et al., Vaccine 2008)
  - Self-controlled case series risk windows (Xu et al., Stat in Med 2010)
- Extending **sequential methods** to observational safety setting
  - Continuous monitoring with maxSPRT (Kulldorff et al., Seq Anal 2011)
  - Accounting for uncertainty in hx controls (Li et al., Stat in Med 2010)
  - Sequential design/analysis challenges (Nelson et al, submitted)
  - Group sequential designs simulation evaluation (Zhao et al., submitted)
- Improving methods for **confounder adjustment**
  - Propensity score stratification (Li et al., accepted at Stat in Biosciences)
  - Regression, generalized estimating equations (Cook et al., submitted)
- Improving method to handle **data complexities**
  - Partially-accrued and missing data (Greene et al., in press at PDS)

VSD: Methods challenges and priorities

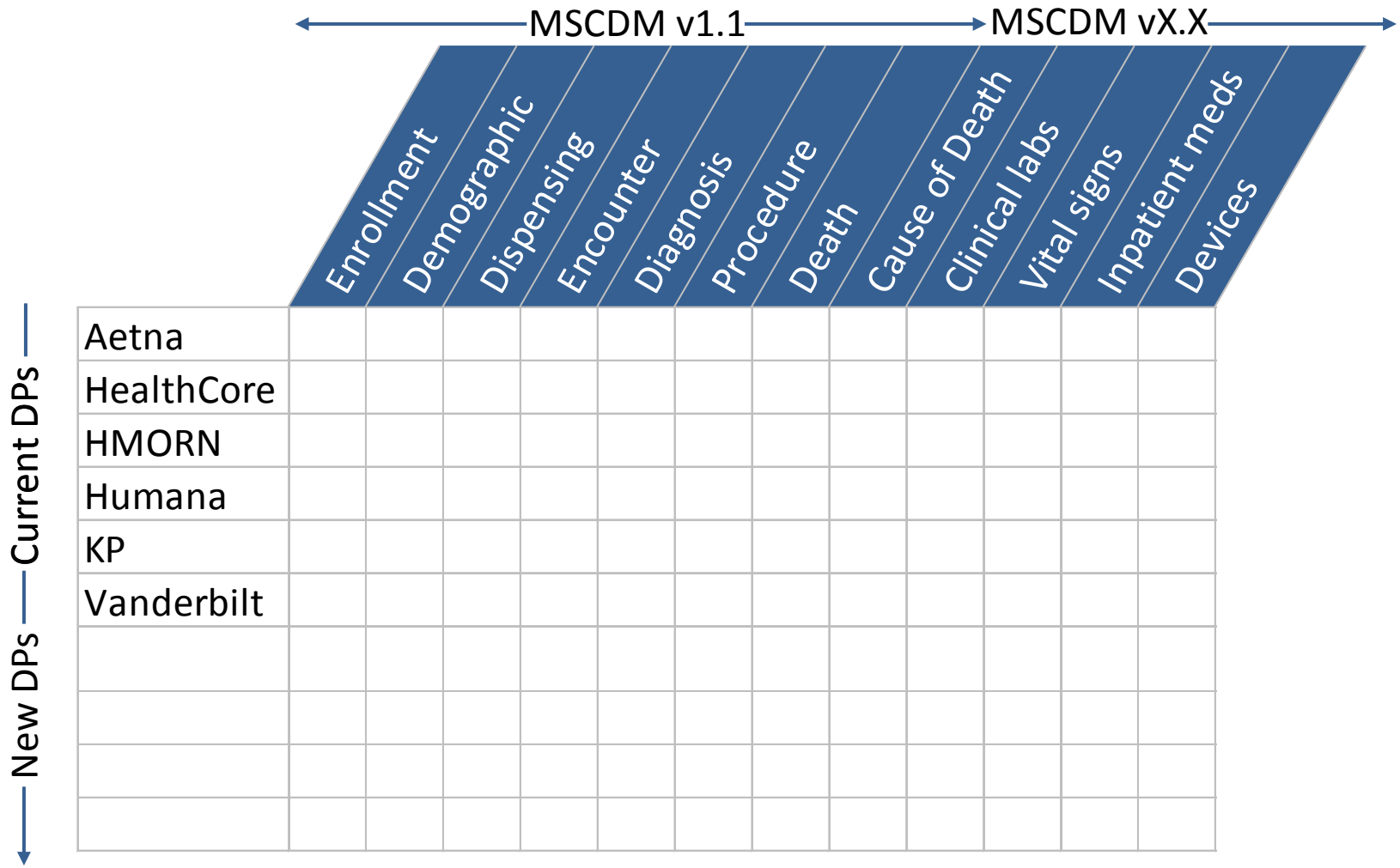
- **Challenges**
  - Optimizing methods given **rare AEs, variable uptake, site heterogeneity**
  - False positive/negative signals due to **misclassification & confounding**
  - Best practices for rapid **signal evaluation** and follow-up
  - Detecting **unanticipated AEs**
- **Priorities**
  - Enhance, customize and evaluate **more sequential approaches**
    - exact tests, delayed starts, risk differences, longer-term outcomes
  - Develop methods to **better account for misclassification & confounding**
    - sequential 2-phase sampling (to get better AE, confounder data)
    - analysis (vs. design) based confounding adjustment, use of propensity scores
  - Solidify system for **rapid signal evaluation**
  - Evaluate and implement **data mining** approaches for signal generation

Map of methodologic domains

Data capacity	Distributed methods	Signal alerting
• **Integrity** – Common data model – Data completeness – Data validity – HOI validation • **Environments** – Claims – EHRs • Ambulatory • Inpatient – Registries – Other (blood banks, genetic data, etc.)	• **Distribution and retrieval** • **Anonymous linkage across sources** • **Distributed multivariable analysis** – Horizontal – Vertical	• **Design & validity** – Expedited design choice – Automated confounding adjustment • **Performance of** – Sequential testing – Non test-based – Decision analytic approaches • **Special aspects** – Drugs, vaccines, biologics, devices

Applications
• Oral antidiabetic agents and MI, rotavirus vaccine and intussusception, etc.

Expansion of MSCDM



Data Capacity

- **Data validation and adjudication**
  - Validation of selected health outcomes of interest (HOI):
    - Myocardial infarction
    - Severe liver injury
    - Anaphylaxis
    - Venous thromboembolism
    - Intussusception
- **Literature reviews on the validity of HOI identification**
  - Review of 20 HOIs completed (accepted at PDS 2011)
  - Review of 20 additional HOI in progress

Map of methodologic domains

Data capacity	Distributed methods	Signal alerting
• **Integrity** – Common data model – Data completeness – Data validity – HOI validation • **Environments** – Claims – EHRs • Ambulatory • Inpatient – Registries – Other (blood banks, genetic data, etc.)	• **Distribution and retrieval** • **Anonymous linkage across sources** • **Distributed multivariable analysis** – Horizontal – Vertical	• **Design & validity** – Expedited design choice – Automated confounding adjustment • **Performance of** – Sequential testing – Non test-based – Decision analytic approaches • **Special aspects** – Drugs, vaccines, biologics, devices

Applications
• Oral antidiabetic agents and MI, rotavirus vaccine and intussusception, etc.

Distributed methods

- Distribution and retrieval
  - Developed increasingly complex query modules
  - Fast turn-around
- Anonymous linkage across sources
  - Of great importance when adding clinically rich data sources to the longitudinal claims data backbone
  - Have identified a candidate method
  - Working group to evaluate such method (starts in June)

Distributed methods

- Evaluating strategies for accessing distributed data (RFTO) (Rassen et al PDS 2010, Med Care 2010, ISPE workshop etc)
  - That allow multivariable confounder balancing
  - That use varying information content to a maximum
  - That allow flexible subgroup analyses
  - Preserve patient privacy
  - Respects plan confidentiality
  - (Velentgas et al. PDS 2008, Rassen et al PDS 2010, Med Care 2010, ISPE workshop etc)
- Provide guidance for MS on valid and practical strategies

Map of methodologic domains

Data capacity	Distributed methods	Signal alerting
• **Integrity** – Common data model – Data completeness – Data validity – HOI validation • **Environments** – Claims – EHRs • Ambulatory • Inpatient – Registries – Other (blood banks, genetic data, etc.)	• **Distribution and retrieval** • **Anonymous linkage across sources** • **Distributed multivariable analysis** – Horizontal – Vertical	• **Design & validity** – Expedited design choice – Automated confounding adjustment • **Performance of** – Sequential testing – Non test-based – Decision analytic approaches • **Special aspects** – Drugs, vaccines, biologics, devices

Applications
• Oral antidiabetic agents and MI, rotavirus vaccine and intussusception, etc.

Design and validity

- Taxonomy project:
  - Expedited choice of design and analytic monitoring approach
  - Identified generic attributes of exposure, outcomes, and relationships
  - Have developed a decision table for fundamental design choice and Year 2 Taxonomy is working on analytic choices
- Self-controlled designs:
  - Came up with clear guidance on (Maclure et al, submitted)
    - Strength/limitations, practicability in a monitoring setting
  - Tested a multivariate SCCS approach (Madigan et al, submitted)

Design and Validity

- Automated covariate adjustment
  - Empirical covariate identification in claims data is essential
    - for improved confounding adjustment
    - for rapid turn-around
  - Empirical approaches have been shown to be superior to investigator identified adjustment in claims
  - Simulation studies have shown that theoretical biases (M-Bias and z-Bias) are not relevant (Myers et al. AJE 2011 in press)
  - A comprehensive approach to automated covariate adjustment is developing for PS and DRS methods (Rassen&Schneeweiss, submitted)

Performance of signal alerting algorithms

- Sequential testing
  - Developed guidance on sequential designs customized for observational safety settings (Nelson et al, submitted)
  - Reviewed methods 'state-of-the-art'
    - Group sequential LRT (inc. maxSPRT, Kulldorff et. al Seq Anal 2011)
    - Conditional sequential sampling procedure (Li et. al Stat Med 2009)
    - Clinical trial group sequential methods (Lan&Demets Stat Med 1994)
    - Estimating equations approach (Cook et al, submitted)
  - Simulation to compare performance (Cook et al, submitted)
    - Type 1 error rate, power, time-to-signal detection
    - Varying outcome prevalence, exposure & confounder complexity
  - Using inverse probability weighting (ongoing Y2 activity)

Performance of signal alerting algorithms

- Non test-based approaches (Pilot)
  - Are available but fair comparisons in a monitoring setting are lacking
  - Pilot work has set up a simulation framework and evaluation statistic (Gagne et al.)
- Safety monitoring and decision science (Pilot)
  - MS recognizes the value of decision analytic approaches in an active surveillance system
  - Pilot work on alerting based on the future value of continued monitoring (Patrick et al.)

Performance of signal alerting algorithms

- Rapid signal validation techniques (WG starting)
  - Develop a framework for follow-up to alerts resulting from signal refinement
    - Data checks
    - Program checks
    - Sensitivity analyses
    - Additional confounding control
    - Endpoint adjudication etc.

Special Aspects

- Vaccines (see VSD)
- Devices
  - Data issues
  - Rapidly changing technologies
  - Exposure-risk window (implantation vs. device itself)
- In-hospital products
  - Data issues
  - Temporality issues
  - Exposure-risk window (end at discharge or later)